Comprehensive database of Manufactured Gas Plant tars – Part C Heterocyclic and hydroxylated PAHs

Authors: Christopher Gallacher*[1], Russell Thomas[2], Richard Lord[1], Robert M. Kalin[1] and Chris Taylor[3]

[1] *Department of Civil and Env. Eng., University of Strathclyde, 75 Montrose St. Glasgow, G1 1XJ UK*

[2] *WSP / Parsons Brinckerhoff, Kings Orchard, 1 Queen St, Bristol, BS2 0HQ UK*

3 National Grid Property, Warwick Technology Park, Warwick, CV34 6DA UK

**Corresponding Author: Christopher Gallacher, christopher.gallacher@strath.ac.uk*

Abstract

RATIONALE

Coal tars are a mixture of both organic and inorganic compounds and were produced as a by-product from the manufactured gas and coke making industries. The product tar compositions varied depending on many factors such as the temperature of production and the type of retort used. For this reason a comprehensive database of the compounds found within different tar types is of value to both understand how their compositions differ and what potential chemical hazards are present. This study focuses on the heterocyclic and hydroxylated compounds present within a database produced from 16 different tars from 5 different production processes.

METHODS

Samples of coal tar were extracted using accelerated solvent extraction (ASE) and derivatized post extraction using *N,O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS). The derivatized samples were analysed using two-dimensional gas chromatography combined with time-of-flight mass spectrometry (GCxGC/TOFMS).

RESULTS

A total of 865 Heterocyclic compounds and 359 hydroxylated polycyclic aromatic hydrocarbons were detected within 16 tar samples produced by 5 different production process. Both the heterocyclic and hydroxylated PAH content varied greatly with the production process used with the heterocyclic compounds giving information about the feedstock used. Of the 359 hydroxylated PAHs detected the majority would not be detected without the use of derivatization.

CONCLUSIONS

Coal tars produced using different production processes and feedstocks produced tars with significantly different heterocyclic and hydroxylated contents. The concentrations of the individual heterocyclic compounds varied greatly even within the different production processes and provided information about the feedstock used to produce the tars. The hydroxylated PAH content of the samples provided important forensic information that would otherwise not be obtained without the use of derivatization and GCxGC/TOFMS.

Keywords: GCxGC/TOFMS, Coal Tar, Environmental Forensics, Derivatization

Introduction

Coal tar is a complex mix of inorganic and organic compounds produced primarily by the former manufactured gas industry and heterocyclic compounds are a major group of compounds of interest present within coal tars. Of particular interest in samples of coal tar, or coal tar contaminated soils, are those containing oxygen, sulfur and nitrogen. The composition of sulfur (PASH), oxygen (PAOH) and nitrogen-PAHs (PANH) in tar largely reflects those present in the parent coal[1], with some temperature-dependent alteration[2]. This makes heterocyclic compounds potential compounds of forensic interest as they may provide information about the parent coal used to produce the tar.

Sulfur containing compounds make up an important group of compounds present within coal tars with heterocyclic sulfur compounds dominating. The sulfur content within coal is present either as inorganic compounds, such as pyrite and other sulfides or organic sulfur compounds, such as poly aromatic sulfur hydrocarbons (PASHs). The organic sulfur content (OSC) of coal is determined by the original organic matter that formed the coal deposits and takes the form of aliphatic and aromatic thiols, sulphides, disulphides and heterocyclic combinations of thiophenes and dibenzothiophenes[3]. PASHs exist in an even greater variety of structures compared to PAHs due to the presence of sulfur within the ring structure, therefore the number of isomers and alkylated isomers can be extremely high making quantification and identification of individual PASHs isomers difficult[4]. The occurrence of PASHs (also known as thiaarenes) in environmental samples and fossil fuels as well as their mutagenic and carcinogenic potential has been reported[4].

Nitrogen is present in all fossil fuels and is associated almost exclusively with the organic portion of the crude material[5]. It usually makes up around 0.5% of crude petroleum but is found in higher concentrations (1-2%) in shale oils and coal[5]. The toxicity of nitrogen aromatic compounds greatly depends on the structure and number of fused rings. Several studies have found that nitrogen-containing substituents, such as nitro- and amino- functional groups can enhance toxicity by up to 100-fold. This means that although the nitrogen content of the parent coal may be low the potential negative health effects from nitrogen containing polycyclic aromatic compounds (PANH) cannot be ignored. While in general PANH compounds are present in lower concentrations than their non-substituted PAH-analogues their higher water solubility leads to a higher potential bioavailability and increased potential toxic effects within the environment[6]. PANHs are highly stable relative to neutral PAH's and can persist through severe thermal conditions and so are possible compounds of toxicological interest[7].

Oxygen containing compounds are also present within coal tar in the form heterocyclic oxygen containing compounds and non-heterocyclic oxygen containing compounds such as aromatic ketones. Oxygen containing compounds can be toxic and mutagenic and are more mobile within the environment than their parent PAHs due to their increases solubility in water[8]. This makes them of special concern in ground water as they can more readily leach from plumes of coal tar DNAPL to contaminate drinking water. Benzofuran and dibenzofuran are two important oxygen containing PAHs that are derived from coal tar and are important biomarkers for coal tar contamination.

Hydroxylated PAHs, sometimes referred to as tar acids or phenolics, are another major group of oxygen containing compounds found within coal tars with the alkyl phenols dominating[9]. The tar acids composition of vertical retort tars varied greatly from other tars with VR tars containing 6-7 times more tar acids than Coke oven tars and 3-4 times more tar acids than horizontal retort tars[10]. The compounds present within the tar acid fraction are also very different with HR and CO tars containing significantly more phenol and cresols than VR tars, with phenol making up 25% of HR tar acids and 18% of CO tar acids[10]. VR tars acids on the other hand contain only 6% phenol, with cresols only making up a small proportion, with high boiling tar acids dominating[10]. A database of the heterocyclic and hydroxylated PAH content of coal tar samples from different production processes has never previously been published.

This study presents the in depth analysis of the heterocyclic and hydroxylated content of 16 tar samples produced by 5 different production processes. A full database of the compounds found within the samples has been published in Gallacher et al., 2017a[11] and an in depth analysis of the aliphatic and aromatic content can be found in Gallacher et al., 2017b[12].

Materials and Methods

Samples:

A total of 16 tars samples, coming from 5 different tar production processes, were extracted and analysed. The different processes and sample numbers are listed below (full site information, and TIC chromatograms, for each sample can be found in the supplementary information):

Low Temperature Horizontal Retort: DNAPL009 (D9L) and DNAPL016 (D16L)
Vertical Retort: DNAPL002-006 (D2-D6V) and DNAPL020 (D20V)
Horizontal Retort: DNAPL007 (D7H), DNAPL008 (D8H), DNAPL010 (D10H) and DNAPL017 (D17H)
Carbureted Water Gas: DNAPL013 (D13C) and DNAPL014 (D14C)
Coke Oven: DNAPL018 (D18CO) and DNAPL019 (D19CO)

Methods:

All solvents used were of analytical grade purchased from Fisher Scientific (Loughborough, U.K.) and D_{10}-phenanthrene, which was used as an injection standard, was purchased from Sigma-Aldrich (Gillingham, U.K.). D_8-naphthalene, D_{10}-fluorene, D_{10}-fluoranthene and D_{10}-pyrene, which were used as recovery standards, was purchased from Sigma-Aldrich (Gillingham, U.K.). *N,O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylchlorosilane (TMCS) was purchased from Sigma-Aldrich (Gillingham, U.K.). Silica was purchased from Fisher Scientific (Loughborough, U.K.)

Extraction was performed using an Accelerated Solvent Extraction system (ASE 350 Dionex, Camberley, UK) using 10 mL stainless steel extraction cells. Approximately 0.5g of tar was mixed with an equal amount of diatomaceous earth in a 1:1 ratio. Prior to extraction the samples were spiked a recovery standard. Extraction cells were lined with 2 Dionex glass fibre filter papers and packed with 3g of silica gel 60 deactivated with 10% water. The sample mixture was then loaded into the cells and any residue

was recovered with excess diatomaceous earth. Dichloromethane was used as the extracting solvent for all extractions. ASE was performed at 100°C and 10 MPa, using one dynamic (7 min) and two static (5 min each) extractions. A flush volume of 150% and purge time of 60 s was used. The extracts were concentrated to 1 mL using a Büchi Syncore Analyst (Oldham, U.K). The extracts were then made up to exactly 10 mL using *n*-hexane. A 1 mL aliquot was then transferred to an auto sampler vial prior to analysis and spiked with D_{10}-phenanthrene. All samples were derivatized using 100ul of BSTFA with 1% TMCS placed in an oven at 70°C for 1 hour.

GCxGC TOFMS analysis was performed using a Leco Pegasus 4D (St. Joseph, Michigan, USA) time of flight mass spectrometer, connected to an Agilent 7890A (Wilmington, Delaware, USA) gas chromatograph equipped with a LECO thermal modulator. The TOF ion source temperature was 200 °C and the mass range 45 and 500u was scanned at a rate of 200 spectra/second. The detector voltage was set at 1700 V with an electron ionisation energy of 70 eV.

All standards and extracts were analysed with the following primary oven temperature programme: 60°C isotherm for 2 minute, then ramp at 10°C/min to 110°C, then ramp at 3°C/min to 310 °C, and isothermal at 310°C for 15 minutes. The secondary oven and modulator temperatures were programmed at a 10 °C offset relative to the primary oven. The modulation period was 6 seconds with a 1.3 second hot pulse time and a cool time of 1.7 seconds. The injection port temperature was set to 250 °C and set to split injection with a split ratio of 50 and an injection volume of 1μL. Helium was used as the carrier gas, with a flow rate of 1.0 mL/min.

The reversed polarity column set that was used comprised of a mid-polarity TR-50 MS supplied by Thermo Scientific (Hertfordshire, U.K.) (30 m × 0.25 mm i.d. × 0.25 μm film thickness) as the primary column and a non-polar Rtx-5SilMS supplied by Thames Restek (Buckinghamshire, U.K.) (1.5 m × 0.25 mm i.d. m × 0.25 μm film thickness) as the secondary column, connected via a Thames Restek Press-tight connector.

The chromatograms from each sample were processed using Leco ChromaTOF software (Version 4.50.8.0) to search for, identify and align all peaks with a signal-to-noise ratio greater than 10. As it would be impractical to purchase standards of every compound present within the samples any concentration comparisons are relative comparisons of the peak area for each compound adjusted for sample weight and recovery. Only direct like for like compound comparison was performed, as the detector response will remain equivalent.

Results and Discussions

Low Temperature Horizontal Retort (LTHR):

Early low temperature horizontal retorts (LTHR) were heated directly by radiant heat from a shallow fuel bed of coke, beneath the furnace. This design only heated the retorts to around 600°C[13] and as a result the amount of gas produced was fairly low and the decomposition of the organic compounds within the tar was limited, leading to tars more similar in nature to the parent coal[13].

Table 1 shows the total number of derivatized, PAHN, PAOH, PASH and mixed heterocyclic compounds detected within each of the tar samples, and also includes

whether the tar is in the form of pure DNAPL or DNAPL contaminated soil or water. The total number of hydroxylated PAHs present within the LTHR tar samples varies greatly with D9L containing 203 and D16L containing only 60 individual compounds. The relative peak areas of select derivatized compounds is shown in table 2 and shows that while sample D16L contains more phenol than D9L the relative concentrations of the remaining hydroxylated PAHs are higher in D9L relative to D16L. The relative difference between concentrations increased with increasing molecular weight with D9L containing significantly more naphthalene-2-ol, 4-hydroxyfluorene and 2-hydroxyfluorene, than D16L. This large difference in concentration is likely down to the fact D9L is in the form of pure DNAPL and D16L is in the form of DNAPL contaminated soil. This suggests that the original tar from which D16L was derived from had significantly higher concentrations of lower molecular weight hydroxylated PAHs relative to D9L, as the difference between the concentration of for example o-cresol is much smaller than that of 4-hydroxyfluorene. The fact that D16L contains more phenol than D9L further reinforces that the original tar contained more lower molecular weight hydroxylated PAHs. This suggests that the phenolic content of LTHR tars can vary widely, and this is likely related to the feedstock coal used during the manufactured gas process. The phenol and cresol content of D16L is higher than any of the other DNAPL contaminated soils and is higher than the concentrations found in the pure DNAPLs D8H, D17H, C13C, D20V and D18CO suggesting the original concentrations of these compounds would have been significant.

Sample D9L contains the most heterocyclic compounds of the LTHR tars with 194 PANHs, 146 PAOHs and 233 PASHs. D16L contains far fewer heterocyclic compounds with only 88 PANHs, 103 PAOHs and 78 PASHs. While sample D9L contains more PANHs, relative to D16L, the relative concentrations of many of the shared PANHs, shown in table 3, are higher in D16L. Sample D16L contains more benzonitrile, quinoline, 1-naphthalenecarbonitirle, carbazole and benzo[h]quinoline, than D9L. This suggests that the nitrogen content of the coal that produced sample D16L was significantly higher than that which produced D9L. Sample D9L also contains the lowest concentrations of PANHs of all the pure DNAPLs analysed. The relative concentrations of select PAOHs is shown in table 4 and clearly demonstrates that with the exceptions of 9,10-anthracenedione and fluoren-9-one, sample D9L has higher relative concentrations of PAOHs, relative to D16L. The relative concentrations of select PASHs is displayed in table 5, it shows that in all cases D9L has significantly higher concentrations of PASH, relative to D16L. Overall this suggests that the original coal used to produce D9L had a significantly higher sulfur content, and possibly a higher oxygen content. While the locations of the sites are confidential it should be noted that they would not have received coal from the same coalfields and so the type of coal used may have differed.

Horizontal Retort (HR):

The development of gas producers. regenerative furnaces and improved retort materials allowed later horizontal retorts (HR) to operate more efficiently and at higher temperatures capable of exceeding 1000^{o}C[14]. This led to greater decomposition of the organic compounds within the tars produced. The large surface area present within horizontal retorts allowed for a greater opportunity for contact of the gases with the heated retort sides and so allowed for a greater degree of thermal cracking[15].

The total number of derivatized compounds in each of the HR samples, shown in table 1, is fairly consistent with D7H containing 132, D8H containing 152, D10H containing 120 and D17H containing 108 individual compounds. The only two samples with more hydroxylated PAHs than the HR tars are D20V and D9L and of all the sample types HR tars have the most consistently high number of individual hydroxylated PAHs. While sample D8H contains the most individual derivatized compounds the relative concentrations of the derivatized compounds, a selection of which are shown in table 2, are higher in sample D7H. Sample D7H contains the highest phenol, cresols, naphthalen-2-ol, 2-hydroxyfluorene and hydroxypyrene concentrations of any sample within the database. High concentrations of phenol and the three cresol isomers are also present within D10H. While sample D8H does not contain the highest concentrations of derivatized compounds the concentrations present are not insignificant with sample D17H containing the least number of individual compounds and lowest, overall, relative concentrations within the HR tars.

High phenolic content is a major characteristic of low temperature coal tars (650^{o}C) and medium temperature coal tars (800^{o}C)[16]. While this is the case for the LTHR tars, which are produced at temperatures around 600^{o}C, our results suggest it also applies to the HR tars which were produced at temperatures in excess of 1000^{o}C suggesting that high phenolic content is not entirely temperature dependent and retort type does play a role. The high amount of oxygen and water vapour available within the horizontal retort may explain the presence of abundant phenolics[17]. It should also be noted the production process used to produce the middle temperature coal tar analysed by Shi et al., 2012[16], the Lurgi process, is significantly different to any of the production processes used to produce any of the samples within the database. The high pressure used in the Lurgi process would likely affect the overall composition of the tar produced differently to other gasification processes.

The heterocyclic compositions of the HR tars varies greatly with the greatest number of heterocyclic compounds found in D8H which contains 594 compounds including 187 PANHs, 133 PAOHs, 263 PASHs and 12 mixed heterocycles. Sample D7H contains 570 heterocyclic compounds including 200 PANHs, 153 PAOHs and 211 PASHs and 6 mixed heterocycles and sample D17H contains 539 heterocyclic compounds including 181 PANHs, 134 PAOHs, 213 PASHs and 11 mixed heterocycles. Sample D10H contains the least number of heterocycles, of both the HR tars and pure DNAPLs, with 374 compounds including 141 PANHs, 80 PAOHs, 143 PASHs and 10 mixed heterocycles.

The PANH composition of the HR tars varies widely with the relative peak areas of select PANHs shown in table 3. Sample D17H contains the lowest relative concentrations of the PANHs shared by the HR tars with few exceptions. Sample D7H contains both the largest number of individual PANHs as well as the highest relative concentrations of the PANHs shared by the sample, although this is not the case for every compound such as benzonitrile and carbazole. While sample D7H does contain higher concentrations of the majority of the shared PANHs the concentrations found within D7H, D8H and D10H are relatively similar, compared to D17H. This suggests that the nitrogen content of the feedstock coal used to produce D17H was significantly different to the other HR tars.

The PAOH composition of the samples varies greatly, with the relative peak areas of select PAOHs shown in table 4, and as with the PANHs sample D17H contains the lowest relative concentrations of the PAOHs that the HR tar samples share. Sample D17H contains significantly lower concentrations of both benzofuran and C_1-benzofuran than the other HR tars. Unlike the PANHs sample D8H contains the highest concentrations of PAOHs, but not the largest number of individual compounds. Sample D8H also contains the highest concentrations of benzofuran and C_1-benzofuran within the database. As a general trend sample D7H contains higher concentrations of PAOHs than D10H. This means the order of concentration goes from sample D8H to D7H to D10H and finally D17H. The concentrations of PAOHs within the HR samples follows the same concentration trend as the EPA16 PAHs[12] with sample D8H containing the highest concentrations and D17H the lowest. This is not the case for the PANHs as sample D8H contains the highest concentrations of PANHs but not the EPA16 PAHs. This suggests that the oxygen content of the coals used to produce the HR tars were similar with the nitrogen content varying to a greater degree.

One compound of note is 9H-fluoren-9-one, which is present in high concentrations in sample D17H relative to the other HR tars. This high concentration is completely at odds with the other PAOH concentrations for the HR tars and this is likely related to the source of the compound. Sample D17H also contains 4H-cyclopenta[def]phenathren-4-one which is not detected within any of the other HR tars. These two compounds are of particular interest as fluorenone can be produced from the metabolism of fluorene[18] and fluoranthene[19] and increased concentrations of fluorenone and 4H-cyclopenta[def]phenathren-4-one have been reported during the degradation of creosote, a coal tar distillate, contaminated soils[20]. Sample D17H also contains significantly more 9,10-anthracenedione than D7H and D8H with 9,10-anthracenedione being completely absent from D10H. Increasing concentrations of fluorenone and 9,10-anthracendione have also been reported during the degradation of coal tar oil[21]. This could suggest that sample D17H may have undergone significant microbial degradation relative to the other HR tars and could suggest evidence for microbial activity within tar DNAPL, although it is also possible these compounds occurred within the feedstock coal or were produced during the gas production process.

When the database is considered as a whole sample D17H contains the highest relative concentrations of 4H-cyclopenta[def]phenathren-4-one and 9,10-anthracenedione but not 9H-fluoren-9-one. Sample D16L contains the highest concentrations of 9H-fluoren-9-one and this is likely down to the nature of the sample as sample D16L is a tarry soil and so microbial breakdown of the tar is more likely to occur than in a pure DNAPL. Significant concentrations of 4H-cyclopenta[def]phenathren-4-one and 9H-fluoren-9-one are also present within sample D20V. Of all the samples in the database D16L, D17H and D20V contain the most 4H-cyclopenta[def]phenathren-4-one, 9,10-anthracenedione and 9H-fluoren-9-one. Both D17H and D20V are in the form of pure DNAPL, with D17H being taken from inside a tar tank and D20V taken from a drain, which suggests that bacterial metabolism may be occurring within the pure DNAPL and DNAPL contaminated soils.

The PASHs content of the samples varies greatly with both the largest number, 263, and highest concentrations found in sample D8H, with the concentrations of select PASHs shown in table 5. Sample D8H also contains the second largest number of PASH compounds in the database with only D14C containing more PASHs, with 280. Sample D17H contains the smallest number of individual PASHs of the HR tars and as with the PANHs and PAOHs the lowest relative concentrations. Both samples D7H and D10H contain similar numbers of individual PASHs with D10H containing higher relative concentrations of lower molecular weight PASHs such as benzenethiol and benzo[b]thiophene, whereas D7H contains relatively higher concentrations of higher molecular weight PASHs such as dibenzothiophene. This suggests that the original parent coal that produced D7H contains more high molecular weight PASHs, whereas D10H contained more lower molecular weight PASHs. Sample D8H may have been produced from a feedstock coal with a significantly higher sulfur content than the other HR tars. It is possible that sample D17H contains relatively less tar than the other samples or the sulfur content of the feedstock coal may have been significantly lower.

Vertical Retort Tars (VR):
Vertical retorts were developed in the first decade of the 20th century, they used a retort vessel which was rotated by 90^{o} and could allow for the gas making process to be run continuously. Vertical retorts were operated as a high temperature process with temperatures achieved similar to those of horizontal retorts, however, because of their design they produced significantly different tars. A temperature gradient existed within a vertical retort, especially in a continuous system where coal was continuously fed down by gravity. The tar fog generated could escape vertically up through the coal bed with minimal exposure to the hot surfaces of the retort, reducing secondary decomposition. Vertical retort tars should contain 6-7 times as much tar acids as coke oven tars and 3-4 times as much as horizontal retort tars[10]. Whilst the tar acid, and therefore phenolic, content is high, the phenol content of the tar acids in VR tars is low making up only 6%, with small amounts of cresol, whereas phenol accounts for 25% of the tar acid content of HR tars and 18% of the tar acid content for CO tars[10].

The total number of derivatized compounds, shown in table 1, within the VR tars varies greatly with sample D2V containing 2; D3V containing 1; D4V containing 75; D5V containing 21; D6V containing 31; and D20V containing 231. Sample D20V contains the most individual hydroxylated PAHs within the database. Sample D20V is in the form of pure DNAPL, taken from a drain, whereas samples D2V TO D6V take the form of DNAPL contaminated soils. This means that the VR tars fall into two distinct groups, with samples D2-D6V also providing an interesting case study as all the samples come from the same site.

While sample D20V contains the most hydroxylated PAHs within the database it is also the only sample in the database in which phenol was not detected. The relative concentrations of derivatized compounds, shown in table 2, show that the overall concentrations of the selected compounds are low within all of the VR tars. The hydroxylated PAH content of D20V however is dominated by alkylated phenols with 133 alkylated phenols between C_1 and C_8 detected. A main characteristic of VR tars is the presence of a large abundance of high boiling phenolic compounds[1], those larger than C_2-phenol[22], and this is clearly seen in D20V with the sample containing

124 phenolic compounds within this range. Of the 124 phenolic compounds present within D20V that boil above C_2-phenol 41 are unique to sample D20V. While D20V contains a large number of individual high boiling phenols it does not contain the highest concentrations of C_3-phenol or C_3-phenol 1DB isomers. However for the alkylated phenols in the C4 range and higher the relative concentrations present with sample D20V are higher than all other samples in the database with the exception of D9L. Sample D9L shares 74 alkylated phenols with sample D20V of which only 13 are present in higher concentrations in sample D20V. This suggests that while a large high boiling phenolic content is indicative of a VR tar[1] it also may apply to LTHR tars. While the literature does state that VR tars should have a higher tar acid content than HR tars[10] this is likely referring to HR tars and not LTHR tars, as these would have been long phased out by the 1950's. When the high boiling phenolic content is compared to the HR tars, which are also all pure DNAPLs, higher concentrations of the majority of the high boiling phenolics are present in D20V, as would be expected within a VR tar. It should also be noted that sample D20V came from a drain on the site it was sampled from and so it is possible that phenol has been lost from the sample due to it's higher water solubility than alkylated phenols. These results suggest that it is very important to distinguish between LTHR and HR tars when considering the potential phenolic content of a sample of tar that could have been produced by either HR or LTHR.

The overall heterocyclic composition of the VR tars varies greatly with sample D20V containing significantly more heterocyclic compounds than the other VR tars with 498 individual heterocycles detected including 220 PASHs, 135 PANHs, 132 PAOHs and 11 mixed heterocycles. Sample D2-D6V come from the same sample site with sample D2V containing 200 individual heterocyclic compounds with 151 PASHs, 17 PANHs, 31 PAOHs and 1 mixed heterocycle. Sample D3V contains 262 individual heterocyclic compounds with 181 PASHs, 21 PANHs, 59 PAOHs and 1 mixed heterocycle. Sample D4V contains 404 individual heterocyclic compounds with 205 PASHs, 94 PANHs, 103 PAOHs and 2 mixed heterocycle. Sample D5V contains 292 individual heterocyclic compounds with 157 PASHs, 60 PANHs, 73 PAOHs and 2 mixed heterocycles. Sample D6V contains 290 individual heterocyclic compounds with 139 PASHs, 67 PANHs, 82 PAOHs and 2 mixed heterocycles. While samples D2-D6V come from the same site the samples do not all come from the same part of the site with sample D2V coming from the opposite side of the site to samples D3-D6V. Samples D3V, D4V and D5V come from the same backfilled tar tank and sample D6V comes from a borehole downhill from the tar tank that samples D3-D5V were taken. While samples D3V, D4V and D5V come from a tar tank they are not in the form of pure tar DNAPL as the tar tank had been backfilled with soil at some point during the sites history.

Sample D20V contains the most PANHs within the VR tars with 135 as well as the significantly higher concentrations than D2V TO D6V, with selected PANH concentrations shown in table 3. Sample D20V also contains the most PAOHs of the VR tars again with significantly higher concentrations relative to D2V TO D6V, with select PAOH concentrations shown in table 4. The total number of PASHs present within the VR tars is relatively high with sample D20V containing the most with 220 and sample D2V containing the least, of the VR tars, with 151 individual PASHs. Of the 220 PASHs present within D20V 50 were not detected within any of the other VR tars. Of the 170 remaining compounds 125 individual PASHs are present within

sample D20V at higher concentrations than any of the other VR tars. A total of 45 PASHs are present in lower concentrations in sample D20V than in at least one of the remaining VR tars suggesting that overall D20V contains both more individual PASHs and higher relative concentrations. The majority of the compounds present in lower concentrations in D20V are alkylated PASH isomers including 3 C_1-benzothiophene isomers; 6 C_2-benzothiophene isomers; 3 C_1-dibenzothiophene isomers; and 10 C_2-dibenzothiophene isomers. Sample D6V is the only sample in the database that lacks C_2-thiophene, although it does contain many alkylated thiophenes above C_2. The differences in PANH, PAOH and PASH concentrations in the VR tars is likely down to the fact D20V is pure DNAPL while the other VR tars are soil contaminated DNAPL. The results suggest that the feedstock coal used to produce D2V TO D6V may have had a higher sulfur content than that used to produce D20V.

Since samples D2V to D6V come from the same sample site they form an interesting case study showing the differences in contaminant compositions within samples from the same site. The samples can be broken down into three groups with D2V taken from a different part of the site than D3V-D6V and D6V taken from a borehole downhill from the tar tank that contained the remaining three samples. While all of the samples have been identified as being produced by the VR process, by both historical site information and statistical analysis[17], there is a significant degree of variation in the number of individual compounds and the relative heterocyclic and hydroxylated PAH contents. Sample D3V contains only a single hydroxylated compound, phenol, at significantly lower concentrations than samples D4V and D6V, but not D2V or D5V. Sample D2V only contains two hydroxylated PAHs in the form of phenol and m-Cresol, although both are only present in minor concentrations. Sample D5V contains more individual compounds than both D2V and D3V but at lower concentrations than those found in D6V and significantly lower concentrations than D4V. The highest number and highest concentrations of hydroxylated PAHs detected within D2V to D6V are present within sample D4V. This is also the case of the PANHs, PASHs and PAOHs with sample D4V containing, for the most part, the highest concentrations of the shared compounds. Since samples D3V, D4V and D5V were taken from the same tar tank it would be expected for them to have the most similar compositions and this is not the case. This suggests that great care should be used when analysing contaminated samples from the same sample site as significantly different tar compositions can be present, although this could also be a function of the relative amounts of tar DNAPL present within the soil samples.

Coke Oven tars:

Coke ovens are the only modern day source of coal tar still in operation within the UK[23]. Two distinct forms of coke oven tars exist, with low temperature coke ovens tars produced at <700^{o}C producing phenolic compounds as well as PANHs[24]. The other form of coke oven tar comes from high temperature coke ovens, which operate in excess of 700^{o}C and in the case of metallurgical coke between 900^{o}C and 1095^{o}C[25], with PANHs being converted into ammonia, hydrogen cyanide, pyridine bases and nitrogen[22]. Both D18CO and D19CO would be considered high temperature coke oven plants as they produced metallurgical coke, however sample D18CO originated from a newer (1970s) and more efficient coke oven plant which operated at a higher temperature than the plant which produced sample C19CO, which was built in the 1930s.

The total number of derivatized compounds detected within the coke oven tars, shown in table 1, is similar with sample D18CO containing 49 and sample D19CO containing 68, the lowest numbers of the pure DNAPLs analysed. The compositions of the coke oven tars are of particular interest as both samples were produced at the same time from the same feedstock material and therefore any major differences in composition are down to the different coke oven temperatures rather than different feedstocks, although minor variations within the feedstock coals will also have an influence. The larger number of individual phenolics present within D19CO is likely down to the lower temperature used, as lower temperatures are associated with higher phenolic content[22]. A total of 43 hydroxylated PAHs are shared between the CO tars and all are present in higher relative concentrations within sample D19CO, with the concentrations of select derivatized compounds shown in table 2.

Both coke oven tars contain a large number of heterocyclic compounds with 50% of the total individual compounds present in D18CO being present in the form of heterocycles and 47% within D19CO. As both samples were produced from the same feedstock coal any difference in the heterocyclic composition is likely down to the difference in the coke ovens themselves, although there may be some variation within the feedstock coal itself. This makes the coke oven tars of particular importance, as the other database samples will have used widely different feedstock coals.

As sample D19CO was produced at a lower temperature than D18CO it may be expected to have higher concentrations of PANHs[24]. This is not the case for the individual number of PANHs present within each sample as D18CO contains 171 PANHs, whereas D19CO contains only 161. In total the samples share 135 PANHs 65 of which are present in higher concentrations in D19CO relative to D18CO. A total of 69 PANHs are present in higher concentrations in D18CO relative to D19CO. The remaining shared compound, indole, is present in essentially the same concentration. As a general trend sample D18CO contains higher concentrations of nitrile containing PANHs and parent PANHs. Sample D19CO contains relatively higher concentrations of alkylated PANHs such as C_1 and C_2 indole; C_1, C_2 and C_3 quinoline, C_1 and C_2 carbazole; and C_1 and C_2 acridine. These results suggest that at higher temperatures the alkylated PANHs are lost in favour of parent PANHs and non-heterocyclic nitrogen compounds such as those containing nitrile. When the database is taken a whole both CO tars contain a significantly higher PANH content than other tar types, with the concentrations of select PANHs shown in table 3.

As samples D18CO and D19CO are both high temperature coke oven tars they would be expected to contain large abundances of pyridines as above 700^{o}C PANHs are converted into, although not exclusively, pyridine bases[24]. While both CO tar samples have the same number of alkylated pyridines, although not the same individual compounds, the concentrations of alkylated pyridines present within D19CO are significantly higher. A total of 4 alkylated pyridines are shared by the CO tar samples with 2 C_2 pyridines, which are present in 6.2 times and 16.1 times higher concentrations in D19CO, and 2 C_3 pyridines, which are present in 1.1 times and 2.4 times higher concentrations in D19CO. These results suggest that the lower operational temperature of D19CO formed pyridines, whereas the higher operational temperature of D18CO may destroy them. This observation suggests that alkyl pyridines are produced at lower temperatures and so would not be a good marker for

metallurgical coke oven operational temperatures. Pyridine itself was not detected within either sample.

The PAOHs follow the same general trends as the PANHS with D18CO dominated by parent PAOHs such as benzofuran and dibenzofuran, shown in table 4, as well as compounds such as phenyl benzofurans and phenyl dibenzofurans. Sample D19CO is dominated by alkylated PAOHs such as C_1 benzofuran, shown in table 4, as well as C_2 and C_3 benzofuran; C_1 and C_2 dibenzofuran; and C_1, C_2 and C_3 benzonaphthofuran. Sample D18CO contains 107 individual PAOHs and D19CO contains 108 with 86 PAOHs shared between the samples. Of the 86 shared compounds 51 are present in higher relative concentrations in D19CO. These results suggest that higher temperature coke ovens result in tars with lower alkylated PAOH contents and higher parent PAOH content. When the database is taken as a whole both CO tars contain significantly higher concentrations of higher molecular weight PAOHs such as dibenzofuran and benzo[b]naphtha[2,3-d]furan, shown in table 4, than the other tar types

The number of PASHs in the two CO tars is similar with sample D18CO containing 128 and sample D19CO containing 133. In total the samples share 105 PASH compounds of which 63 are present in higher concentrations in D19CO relative to D18CO. As with the PANHs and PAOHs sample D18CO is dominated by parent PASHs such as benzo[b]thiophene, dibenzothiophene, shown in table 5, and phenanthro[2,1-b]thiophene . Sample D18CO also contains relatively more phenyl thiophene, phenyl benzothiophene, benzodithiophene and phenanthrothiophene. Sample D19CO contains relative more C_2-thiophene; C_1, C_2 and C_3 benzothiophene; and C_1, C_2 and C_3 dibenzothiophene. This suggests that, as with the other heterocyclic compounds, the higher temperature increases the relative abundances of parent PASHs while decreasing the relative abundance of alkylated isomers. The two CO tars are also the only samples within the database that lack thiophenes larger than C_2. Unlike with the PANHs and PAOHS the PASH content of the CO tars is not the highest in the database.

Carbureted Water Gas (CWG):

The Carbureted water gas process is another gas production process which used both coal, usually in the form of coke, and oil to manufacture gas. It was developed from the water gas process with the addition of oil being sprayed into the carburetor in order to enrich the calorific value of the gas produced. This means that the organic compounds present within the tar produced come primarily from the oil used in the process, as the water gas process alone did not produce significant amounts of tar (when coke was used as the solid feedstock)[24]. This results in a chemical signature that is significantly different to tars derived for coal. For example oil, generally, has a lower nitrogen content than coal with only <0.5% of the content of crude petroleum being present in the form of PANHs, whereas this number goes up to between 1% and 2% in coals[5].

D13C contains only 7 hydroxylated PAHs whereas D14C contains 94. The relative concentrations of the 7 compounds hydroxylated PAHs present in both samples varies greatly with D14C containing significantly more phenol, o-biphenylol and 4-hydroxyfluorene, shown in table 2. The relatively low number, and concentrations, of hydroxylated PAHs present in D13C is likely down to the fact the sample came from

a groundwater contamination plume, whereas D14C came from a tar storage tank. The size of the relative difference between the individual shared compounds suggests that the original D13C tar had a significantly different hydroxylated PAH composition to that of D14C. While sample D14C contains significantly more phenol than sample D13C when the database is considered as a whole the concentration is relatively small and with the exception of D20V all the other pure DNAPLs contain more phenol and cresols than D14C. This suggests that, overall, CWG tars have a lower phenolic content than coal gas and coke oven tars.

The heterocyclic composition of the two CWG samples varies greatly with D13C containing 48 PANHs, 100 PAOHs and 251 PASHs and 2 mixed heterocycles. D14C contains 102 PANHs, 105 PAOHs and 280 PASHs and 2 mixed heterocycles. The relative compositions of the Heterocyclic compounds are interesting as it gives important forensic information about the samples, such as information about the feedstock used. The PANH content of both samples is low with sample D14C containing significantly higher concentrations of the 32 PANHs that the CWG samples share. For example sample D14C contains significantly more carbazole and benzo[c]carbazole, shown in table 3. Overall the PANH content of D14C is lower than the other pure DNAPLs, with the exception of sample D9L, and this would be expected from an oil based tar.

The total number of PAOHs within both samples is similar, however the concentrations of PAOHs are vastly different with sample D14C containing higher concentrations of 57 of the 71 PAOHs the samples share. Despite the fact sample D13C likely contains significantly less tar relative to D14C, as the sample comes from groundwater, 14 PAOHs are present in higher concentrations in D13C relative to D14C. With the exemption of a single compound, a naphthofuran isomer, the PAOHs found in higher concentrations in D13C over D14C are alkylated benzofuran between C_1 and C_4. This suggests that the original tar that sample D13C came from containing significantly higher concentrations of alkylated benzofurans, which may suggest differences in the production process or could be down to the feedstock oil used. In spite of the fact that D13C comes from a contaminated water plume the concentrations of Benzofuran and dibenzofuran are not insignificant suggesting that the original tar may have had a significantly higher heterocyclic oxygen content than that which produced sample D14C.

Both CWG tars contains a large number of PASHs with 251 in D13C and 280 in D14C. Sample D14C contains the most PASHs in the database, with only D8H (263) containing more than D13C. Of the 198 PASHs that the CWG tars share only a single compound is present in higher concentrations in D13C relative to D14C, a C_3-Thiophene 1DB isomer, with all other compounds being present in higher concentrations in D14C, a selection of which are shown in table 5. When the database is considered as a whole sample D14C, which is more indicative of a pure CWG tar DNAPL, contains significant concentrations of PASHs with the highest concentrations of benzo[b]thiophene, dibenzothiophene and C_1-benzothiophene isomers. Despite the fact D13C is not a pure DNAPL and is in the form of DNAPL contaminated groundwater the concentrations of PASHs detected are higher than those present within some pure DNAPLs with for example more benzo[b]thiophene than D20V and D10H and more C_1-benzothiophene than D20V, D7H, D10H, D17H, D9L, D18CO and C19CO.

Conclusion

Coal tars that are produced using different feedstocks and production processes produced tars with significantly different heterocyclic and hydroxylated contents. In total 865 heterocyclic and 359 hydroxylated PAHs were detected. The phenolic content of coal tars varies greatly depending on the production process used to produce the tars and the phenolic content of tars can provide important forensic information. The presence or absence of phenolic compounds within coal tar samples can also potentially provide information on the environmental degradation of the samples, for example if the production process is known to produce a high phenol content but a high phenol content is not detected this may suggest the phenol has been lost to the environment. This is of particular importance when assessing samples that have been exposed to water as phenol is the most water soluble of the phenolic compounds with increased molecular weight decreasing solubility. The presence of high boiling phenolics, those boiling above C_2-Phenol, can also be used to suggest what retort type was used to produce a specific tar. The data produced by the two coke oven tars is of particular forensic interest as both tars were produced from the same feedstock coal but operated at different temperatures. This demonstrates the effect temperature has on the phenolic content of tars produced without the effect of different retort types. The results clearly showed that the lower temperature coke oven tar contained significantly more individual hydroxylated phenolics as well as higher individual concentrations. This suggests that the higher the carbonizing temperature, the greater the decomposition of the phenolic compounds within the tar.

The heterocyclic content of coal tars is also important as it can provide information on the feedstock used to produce the tars. The concentrations of individual heterocylic compounds varied greatly, even within tars produced by the same production process with the coke oven tar sample data again being of particular importance. The information provided by these two samples can be used to suggest what affect temperature has on the tars produced. At temperatures greater than 700^{o}C PANHs are converted into, although not exclusively, pyridine bases and so it would be expected that the higher temperature coke oven tar would contain higher concentrations of pyridines. This however is not the case and the low temperature coke oven contains significantly higher concentrations of the shared alkylated pyridines. It should be noted that pyridine itself was not detected within either coke oven tar sample and the concentration data suggest the higher temperature resulted in the destruction rather than production of alkyl pyridines. It is also expected that lower temperature coke ovens contain higher concentrations of PANHs and of the 135 PANHs shared by the two coke oven tars 65 were present in higher concentrations in the higher temperature oven with 69 present at higher concentrations in the lower temperature oven, with 1 equal compound. The results suggest that the statement that lower temperature coke oven tars contain higher concentrations of PANHs is not strictly true and a more accurate statement is that higher temperatures result in the loss of alkylated PANHs in favour of parent PANHs and non-heterocyclic nitrogen compounds such as those containing nitrile groups. This is also the case for PAOHs and PASHs with higher temperatures resulting in the loss of alkylated PAOHs with higher parent PAOH content.

The analysis was also able to detect several compounds that might be of interest for suggesting biological activity within the tar samples. Fluoren-9-one may be of

particular interest as it can be produced by the metabolism of both fluorene and fluoranthene and was found in sample D17H in concentrations far in excess of comparable tars samples. The concentrations are completely at odds with the PAOH concentrations for the other HR tars. Sample D17H also contains 4H-cyclopenta[def]phenathren-4-one, which was not detected within any of the other HR tars and has been reported during the degradation of coal tar distillate within contaminated soils. This suggests that significant biological degradation of sample D17H may have occurred and this could provide evidence for microbial activity within tar DNAPL as the sample itself is pure DNAPL taken from a tar tank. Biological activity may also have affected sample D16L as this contains the highest concentrations of Fluoren-9-one, however as the sample is tar within soil then biological degradation is not as usual as that present within sample D17H. Biological degradation of D20V may also have occurred, which is another pure DNAPL.

References

1. A. D. McNeil,. Some Notes on the Chemical Composition of Coal-tar Creosote. *The Gas World*, **1952**, 136, 105-108
2. C. Gauchotte-Lindsay, P. Richards, L. A. McGregor, R. Thomas, R. M. Kalin. A one-step method for priority compounds of concern in tar from former industrial sites: trimethylsilyl derivatization with comprehensive two-dimensional gas chromatography. *J. Chromatogr. A*, **2012**, 1253, 154–63.
3. A. R. Diez, A. I. Gonzalez, R. Menhdez, R. S. Moinelo, J. Bermejo. Characterization of coal tars produced under different carbonization conditions by FT-i.r. spectroscopy and extrography. *Fuel*. **1994**, 37, 139-142
4. S. G. Mössner, S. A. Wise. Determination of polycyclic aromatic sulfur heterocycles in fossil fuel-related samples. *Anal. Chem.* **1999**, 71, 58–69.
5. P. Burchill, A. Herod, P. Ernest. Investigation of nitrogen compounds in coal tar products. 2. Basic fractions. *Fuel*. **1983**, 62, 20–29.
6. J. Neuwoehner, A. K. Reineke, J. Hollender, A. Eisentraeger. Ecotoxicity of quinoline and hydroxylated derivatives and their occurrence in groundwater of a tar-contaminated field site. *Ecotoxicol. Enviro. Saf.* **2009**, *72*, 819–27.
7. L. E. Yu, L. M. Hildemann, S. Niksa. Characteristics of nitrogen-containing aromatic compounds in coal tars during secondary pyrolysis. *Fuel*. **1999**, 78, 377–385.
8. S. Lundstedt, P. A. White, C. L. Lemieux, K. D. Lynes, I. B. Lambert, L. Öberg, O. Lars. Sources, Fate , and Toxic Hazards of Oxygenated Polycyclic Aromatic Hydrocarbons (PAHs) at PAH- contaminated Sites. *Royal Swedish Academy of Sciences*. **2007**, *36*, 475–485.
9. Q. Shi, Y. Yan, X Wu, S. Li, K. H Chung, S. Zhao, C. Xu. Identification of Dihydroxy Aromatic Compounds in a Low-Temperature Pyrolysis Coal Tar by Gas Chromatography−Mass Spectrometry (GC−MS) and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS). *Energy Fuels*. **2010**, 24, 5533–5538.
10. R. Forsdike. The Composition of Tar of Various Origins. *The Gas World*, **1952**, 136, 350.
11. C. Gallacher, R. Thomas, R. Lord, R. M. Kalin, C. Taylor. Comprehensive database of Manufactured Gas Plant tars – Part A Database, *Rapid Commun. Mass Spectrom.* **2017**, In Press.

12. C. Gallacher, R. Thomas, R. Lord, R. M. Kalin, C. Taylor. Comprehensive database of Manufactured Gas Plant tars – Part B Aliphatic and Aromatic compounds, *Rapid Commun. Mass Spectrom.* **2017**, In Press.
13. S. M. Harkins, R. S. Truesdale, R. Hill, P. Hoffman, S. Winters. *U.S. Production of Manufactured Gases: Assessment of Past Disposal Practices.* Prepared by Research Triangle Institute. Prepared for Hazardous Waste Engineering Research Laboratory, U.S. Environmental Protection Agency, Cincinnati, OH; EPA 600-2-88-012. **1988**.
14. W. J. Butterfield. *The Chemistry of Gas Manufacture: A Practical handbook of the production, purification, and testing of illuminating and fuel gas, and on the bye-products of gas manufacture*. Edition 3, Volume 1 Materials and Processes London: Charles Griffin and Company, Limited; Exeter Street, Strand, **1904**.
15. S. Young. *Distillation Principles and Processes*. MacMillan and Company, London, **1922**.
16. Q. Shi, N. Pan, H. Long, D. Cui, X. Guo, Y. Long, C.S. Hsu. Characterization of Middle-Temperature Gasification Coal Tar. Part 3 : Molecular Composition of Acidic Compounds. *Energy Fuels*. **2012**, *27*, 108–117.
17. L. A. McGregor, C. Gauchotte-Lindsay, N. Nic Daéid, R.Thomas, R. M Kalin. Multivariate statistical methods for the environmental forensic classification of coal tars from former manufactured gas plants. *Environ. Sci. Tech.* **2012**, 46, 3744–52.
18. M. Grifoll, M. Casellas, J. M. Bayona, A. M. Solanas. Isolation and characterization of a fluorene-degrading bacterium: Identification of ring oxidation and ring fission products. *Appl. Environ. Microbiol.* **1992**, *58*, 2910–2917.
19. I. Kelley, J. P. Freeman, F. E. Evans, C. E. Cerniglia. Identification of metabolites from the degradation of fluoranthene by Mycobacterium sp. strain PYR-1. *Appl. Environ. Microbiol.* **1993**, *59*, 800–806.
20. M. Eriksson, G. Dalhammar, A. K. Borg-Karlson. Biological degradation of selected hydrocarbons in an old PAH/creosote contaminated soil from a gas work site. *Appl. Microbiol. Biotechnol.* **2000**, *53*, 619–626.
21. H. Wischmann, H. Steinhart. The formation of PAH oxidation products in soils and soil/compost mixtures. *Chemosphere*. **1997**, 35, 1681–1698.
22. C. Woolfolk, C. Golumbic, R. A. Friedel, M. Orchin, H.H. Storch. *Charaterization of Tar Acids from Coal-Hydrogenation Oils*, Bureau of Mines Bulletin 487, **1950**.
23. R.A.P Thomas. *The History and Operation of Gasworks (Manufactured Gas Plants) in Britain*, CL:AIRE, in press, **2014**
24. M. J. Hamper. Manufactured Gas History and Processes. *Environ. Forensics*. **2006**, 7, 55–64.
25. W. T. Lankford, N. L. Samways, R. F. Crave, H. E. McGannon. *The making, shaping and treating of steel (10th ed.)*, Association of Iron and Steel Engineers. Pittsburgh, PA: Herbrick & Held. **1985**.

Acknowledgements
We thank the Scottish Funding Council (SFC) Glasgow Research Partnership in Engineering, the University of Strathclyde, WSP Parsons Brinckerhoff and National Grid property for funding support. The authors declare no competing financial interests.

Sample	Derivatized	PANH	PAOH	PASH	Mixed Het	Tar Type	Retort Type
D2V	2	151	17	31	1	SD	VR
D3V	1	181	21	59	1	SD	VR
D4V	75	205	94	103	2	SD	VR
D5V	21	157	60	73	2	SD	VR
D6V	31	139	67	82	2	SD	VR
D20V	231	220	135	132	11	D	VR
D7H	132	200	153	211	6	D	HR
D8H	152	187	133	263	12	D	HR
D10H	120	141	80	143	10	D	HR
D17H	108	181	134	213	11	D	HR
D9L	203	194	146	233	5	D	LTHR
D16L	60	88	103	78	4	SD	LTHR
D13C	7	48	100	251	2	WD	CWG
D14C	94	102	105	280	2	D	CWG
D18CO	49	171	107	128	8	D	CO
D19CO	68	135	108	133	6	D	CO

Table 1: Total individual derivatized and heterocyclic compounds including tar type. SD = DNAPL contaminated soil, D = pure tar DNAPL, WD = DNAPL contaminated water. VR – Vertical Retort, HR = Horizontal Retort, LTHR = Low Temperature Horizontal Retort, CWG = Carbureted Water Gas, CO = Coke Oven.

Compound	D2V	D3V	D4V	D5V	D6V	D20V*	D7H*	D8H*	D10H*	D17H*	D9L*	D16L	D13C	D14C*	D18CO*	C19CO*
Phenol	4	3	48	3	17	ND	16437	2939	8748	2981	2762	5407	2	426	2555	6454
o-Cresol	ND	ND	80	ND	5	2	8544	2593	7456	1818	5316	2447	ND	941	641	2242
m-Cresol	2	ND	130	ND	6	ND	13756	3135	8824	2559	5241	3847	ND	852	725	4565
p-Cresol	ND	ND	22	1	12	2	9714	1424	5690	1689	4014	2395	ND	654	1607	2126
Naphthalen-2-ol	ND	ND	21	7	4	81	546	195	135	25	379	13	ND	41	27	113
o-Biphenyol	ND	ND	3	ND	11	47	75	52	52	28	74	9	3	29	74	85
4-Hydroxyfluorene	ND	ND	18	8	11	113	114	76	46	39	124	4	3	24	20	68
2-Hydroxyfluroene	ND	ND	14	ND	13	106	172	98	51	74	135	7	ND	25	35	86
Anthrol isomer	ND	ND	6	ND	4	60	56	21	14	24	44	10	ND	9	ND	9
Hydroxypyrene isomer	ND	ND	6	ND	ND	82	91	44	38	14	46	7	ND	ND	22	34

Table 2 – Select derivatized compounds peak areas (M*) adjusted for sample weight and internal standard response (10^{-3}) ND = Not detected * = Pure DNAPL

Compound	D2V	D3V	D4V	D5V	D6V	D20V	D7H	D8H	D10H	D17H	D9L	D16L	D13C	D14C	D18CO	C19CO
Pyridine	ND	ND	ND	ND	ND	ND	1017	ND	2421	ND	19	ND	ND	ND	ND	ND
Aniline	ND	ND	ND	ND	ND	ND	264	50	461	ND	ND	ND	ND	ND	128	112
Benzonitrile	ND	ND	ND	ND	ND	ND	4822	4333	7359	682	32	139	ND	ND	5019	2997
Indole	ND	ND	53	ND	ND	7	5656	4924	5021	441	582	ND	ND	404	8562	8562
Quinoline	ND	ND	68	ND	ND	62	9421	6949	7249	3264	339	863	ND	56	13977	11815
1-Naphthalenecarbonitrile	8	8	120	26	14	1725	3439	3370	2001	1650	135	288	333	625	4153	2999
Carbazole	99	95	922	398	376	9731	19612	22092	15952	10196	841	936	605	4282	36879	29046
Benzo[h]quinoline	ND	4	89	40	51	1422	2660	2617	1524	1263	98	222	90	397	4111	3218
Benzo[def]carbazole	7	5	88	33	41	919	1674	1937	1467	856	72	43	25	321	4477	3282
Benzo[c]carbazole	16	ND	88	50	74	1023	1746	1774	1418	940	81	82	25	368	3626	3083
Benzo[b]carbazole	5	ND	93	20	61	696	1204	1125	ND	632	58	57	17	232	2241	2000
Dibenzo(a,g)carbazole	ND	ND	4	ND	ND	26	43	84	63	55	5	ND	ND	25	63	36
Naphtho[2,3-b]carbazole	ND	ND	2	ND	ND	15	25	39	13	20	1	ND	ND	4	50	49
Benzo[def]naphtho[2,3-b]carbazole	ND	ND	6	ND	ND	ND	ND	113	108	19	4	ND	ND	15	313	9
C1-Carbazole a	30	31	83	33	33	522	701	758	<LOQ	292	445	64	64	390	881	1070
C1-Carbazole b	14	12	91	56	64	1027	1516	1576	<LOQ	781	887	67	130	534	1611	1860

Table 3 – Select PANH peak areas (M*) adjusted for sample weight and internal standard response (10^{-3})

Compound	D2V	D3V	D4V	D5V	D6V	D20V	D7H	D8H	D10H	D17H	D9L	D16L	D13C	D14C	D18CO	C19CO
Benzofuran	15	18	712	379	19	1479	13076	22701	16358	2552	6813	1178	3367	4704	4292	2637
C1-Benzofuran	293	207	855	448	39	2976	10138	18755	10399	2781	6269	1783	4238	4176	1655	2611
Dibenzofuran	1101	1191	5459	2444	1904	76003	71868	90038	51369	35441	33061	6643	20118	27346	143262	114602
Benzo[b]naphtho[2,3-d]furan	82	107	477	198	267	6185	8316	8753	4795	3598	3279	808	386	1571	15736	12045
Benzo[b]naphtho[2,1-d]furan	28	23	92	85	101	1575	2568	2707	835	1080	1117	292	55	556	3710	3375
Benzo[b]naphtho[1,2-d]furan	28	15	241	98	128	2924	3904	3959	2168	1714	1709	444	73	788	6090	4558
7H-Benz[de]antracen-7-one	151	12	26	19	12	42	185	151	92	119	32	20	111	142	999	782
9H-Fluoren-9-one	45	39	19	ND	10	374	59	42	ND	622	38	1080	ND	118	113	160
Cyclopenta(def)phenanthrenone	15	12	5	ND	10	78	ND	ND	ND	154	ND	ND	ND	27	ND	ND
9,10-Anthracenedione	66	89	ND	ND	ND	114	25	21	ND	395	10	57	13	76	ND	13
Acetophenone	13	ND	ND	ND	ND	ND	9	24	34	4	ND	ND	17	39	ND	ND

Table 4 – Select PAOH peak areas (M*) adjusted for sample weight and internal standard response (10^{-3})

Compound	D2V	D3V	D4V	D5V	D6V	D20V	D7H	D8H	D10H	D17H	D9L	D16L	D13C	D14C	D18CO	C19CO
Benzenethiol	ND	ND	16	6	ND	0	16	117	36	6	11	ND	17	43	5	ND
Benzo[b]thiophene	4532	5482	11341	4825	705	24701	53079	117656	72839	22981	34459	3060	28399	139923	52444	35579
C1-Benzothiophene a	4028	4617	3185	1270	563	2453	3423	17458	2936	2906	2528	351	6721	49265	1476	1531
C1-Benzothiophene b	3179	1950	1648	628	221	912	2331	7314	2490	1958	2018	291	3288	12752	1609	1563
C1-Benzothiophene c	3293	3776	2340	945	515	973	2977	12948	1562	2059	1504	225	5472	40372	670	761
Dibenzothiophene	5377	5467	5724	2508	1808	11214	10215	29385	9386	8414	5574	476	9494	83734	31716	20930
Naphtho[1,2-b]thiophene	817	864	1041	433	375	7919	6789	9418	4318	3685	3404	457	1518	9664	5843	3830
Naphtho[2,1-b]thiophene	503	517	624	222	197	4791	4464	6027	2855	2373	2244	333	872	6080	3882	2883
Naphtho[1,2-c] or [2,3-c]thiophene	248	254	323	129	127	2065	2340	2952	1390	1186	1235	179	353	2879	1694	971

Table 5 – Select PASH peak areas (M*) adjusted for sample weight and internal standard response (10^{-3})

Supplementary information: Site History

All samples were taken from sample sites located within the UK and stored at 4^{o}C in the dark prior to analysis.

DNAPL001-006 (D2-D6) samples were taken from the same site that manufactured gas between 1836 and 1971. The site had manufactured gas from a wide range of process from low temperature horizontal retorts, high temperature horizontal retort and later vertical retort. D2 was taken from a borehole near a gasholder on the southern boundary of the site. D3, D4 and D5 were taken directly from a former tar tank on the opposite side of the site. This tar tank had been built at the same time as the vertical retort plant had been constructed. D6 was taken from a borehole located near to the tar tank containing samples D3, D4 and D5. All of the samples were produced by the vertical retort process. D2, D3 and D6 are Sandy soil samples heavily contaminated with tar. D4 and 005 are both waterlogged soils that are heavily contaminated with tar. The samples were taken into 250ml glass jars in 2009.

The site initially operated simple horizontal retorts until the gasworks it was redeveloped in 1878 and expanded in 1912 to include five continuous vertical retort beds. This was further extended in the 1920's with the addition of an additional five continuous vertical retorts. From the early 1930's until 1952 the plant remained at capacity, with various improvements to increase efficiency, such as the addition of two water gas plants. The site was expanded again in 1952 with the construction of a purification plant, additional vertical retorts and the development of a storage area for primary flash distillate (a type of light petroleum distillate). A CWG plant and microsimplex oil gasification plant operated on the site and so there is the potential for traces of CWG and oil gas tar within the samples.

DNAPL007 (D7) was taken from a site that operated between 1856 and 1969 and initially produced low temperature horizontal retort gas. However, this later switched to higher temperature horizontal retort gas and the tar sample came from this process rather than the earlier LTHR process. The sample is in the form of pure DNAPL taken at the base of an underground gasholder tank during a remediation project in 2009 and stored within a 500ml glass bottle.

DNAPL008 (D8) was collected from a FMGP, which operated between 1856 and 1971 and may contain traces of CWG tar as the site produced CWG in the later stages of its operation. The majority of the tar would have been produced by the HR process and this was shown in the statistical analysis of the tars (McGregor et al., 2012). The sample is in the form of pure DNAPL in water. The tar sample was taken from a former 250,000 gallon tank associated with the horizontal retort house and was taken from a skimmer pump, which was being used to empty the tar tanks during a remediation process. The sample was taken in 2009 and stored within a 250ml glass bottle.

DNAPL009 (D9) was obtained from a site that ceased production in 1953, with an unknown initial operation date, the site was manufacturing gas using a vertical retorts when it closed, but could also have produced horizontal retort, likely a LTHR, or carbureted water gas. The CWG plant was commissioned in 1902, with vertical retorts being constructed in 1913 and 1931. Statistical analysis showed the samples to have been derived from the LTHR process (McGregor et al., 2012). The sample is in

the form of pure DNAPL in water and was taken from inside a former tar well during a remediation process. The sample was taken in 2010 and stored within a 250ml glass bottle.

DNAPL010 (D10) was obtained from a FMGP, which operated between 1849 and 1981 and mainly operated horizontal retorts, originally using LTHR and later switching to producers fired HR process circa 1890's. The tar samples were collected from a circular concrete tar tank known to be associated with the later HR process. The sample is in the form of pure DNAPL in water. The site also operated a CWG plant so it is possible that CWG tar may also be present within the sample. The sample was taken in 2010 and stored in a 500ml glass bottle.

DNAPL013 (D13) operated from 1885 until the 1970's, the site manufactured gas using horizontal retort until 1939, taking a bulk gas supply from another source. It did continue to maintain and use its CWG plan to supply peak demand. . The sample of liquid NAPL was removed from a groundwater NAPL plume beneath the site by a skimmer pump. The sample was taken in 2010 and stored in a 250ml glass bottle.

DNAPL014 (D14) was taken from the same site as DNAPL017 and was taken from an infilled former gasholder and later a tar emulsion storage tank, used for the storage of NAPL emulsions produced by the CWG process and was present in the form of pure DNAPL. This sample was obtained from a remediation system, which used total fluids pumping. The sample was taken in 2010 and stored in a 250ml glass bottle.

DNAPL016 (D16) was obtained from a FMGP which operated between 1854 and 1946 and was a typical small country town FMGP site. The site is known to have operated hand charged directly fired horizontal retorts which would have produced a tar with low temperature characteristics (LTHR), which was shown in the statistical analysis (McGregor et al., 2012). The sample was recovered from the core barrel of a pioneer rotary drilling rig and is in the form of a tar saturated soil. The sample was taken in 2010 and stored in a 500ml glass bottle.

DNAP017 (D17) and was taken from the same site as DNAPL014 (D14) from a site which operated from 1868 to 1976. The site operated HR, Coke ovens, CWG, gas reforming plant and a by-products works. The samples were obtained from an area of the site where both the CWG and HR had operated. D17 was taken from a tar tank that would have been associated with the HR process. The sample was taken in 2011 and stored in a 500ml glass bottle.

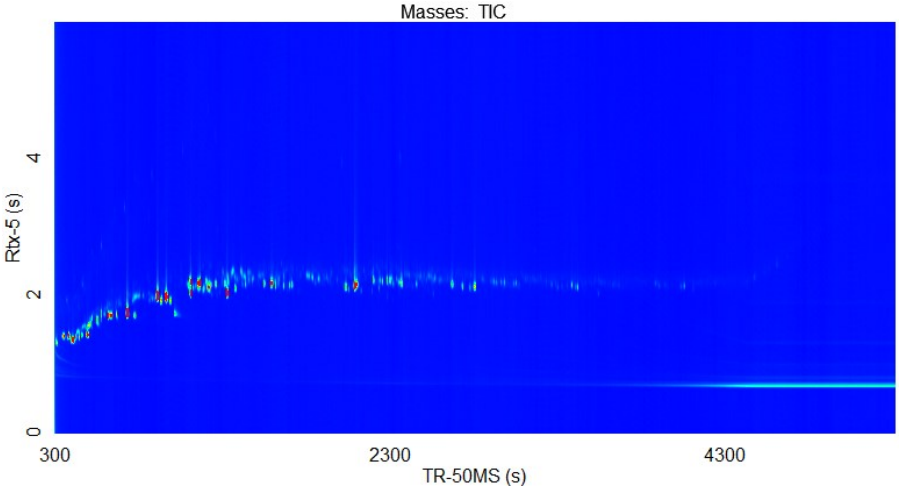
DNAPL018 (D18) and DNAPL019 (D19) were taken from coke ovens that were still in operation at a steelworks. Both samples were taken on the day of coke production with D18 coming from an oven that has been in operation since the 1970's. D19 was taken from a coke oven that has been in operation since the 1930's. Both samples are in the form of pure DNAPL. The coals used in the manufacture of these tars was a blend of up to 12 non British imported crushed coals. The samples were taken in 2011 and stored in 500ml glass bottles.

DNAPL020 (D20) operated from 1885 and ceased production in 1975. The site initially used horizontal retorts until the 1920s before then operating a Vertical retort plant and CWG plant. The coal tar sample was derived from the VR process. The

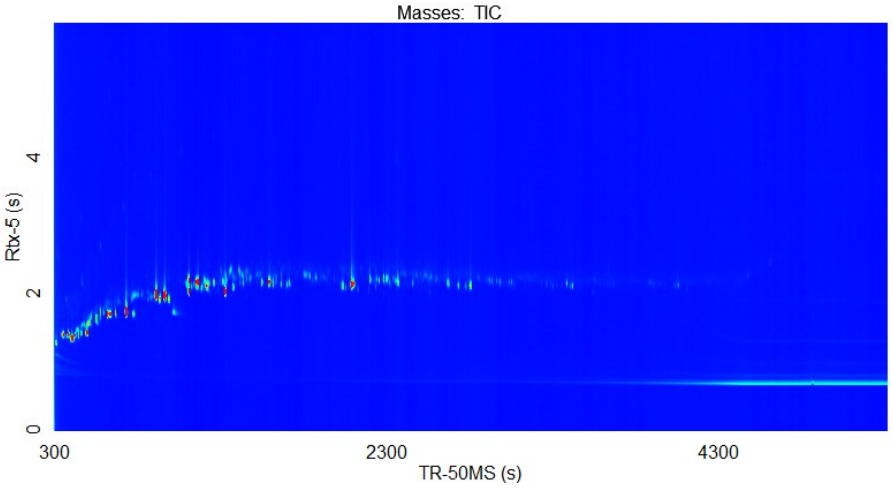
sample is in the form of pure DNAPL taken from a site drain. The site also operated a CWG plant so CWG tar may be present in trace amounts. The sample was taken in 2011 and stored in a 250ml glass bottle.

Sample TIC chromatographs:

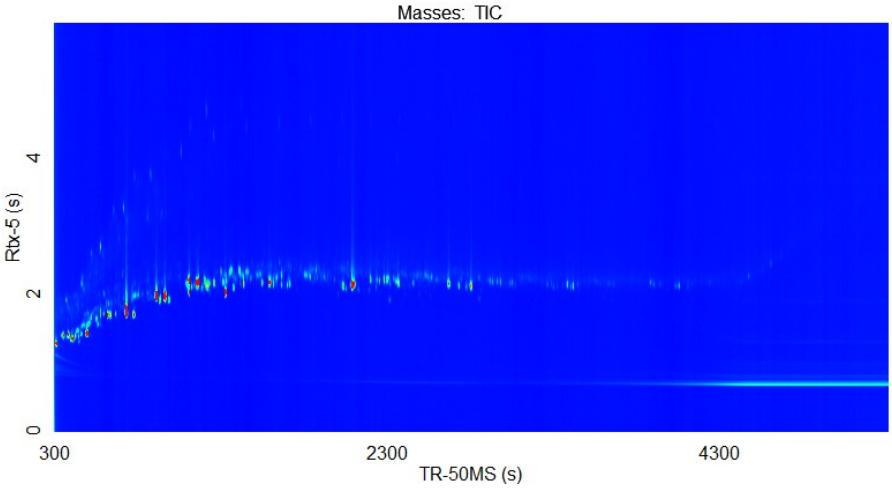
D2V: TIC



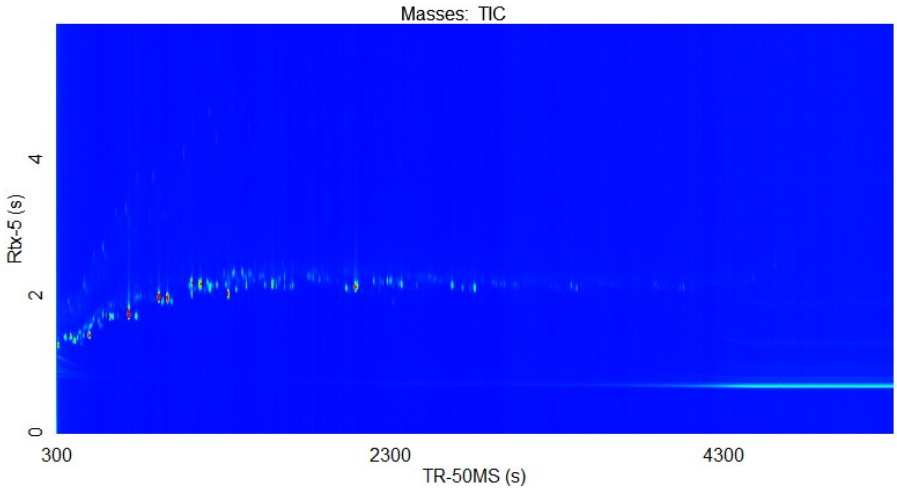
D3V: TIC



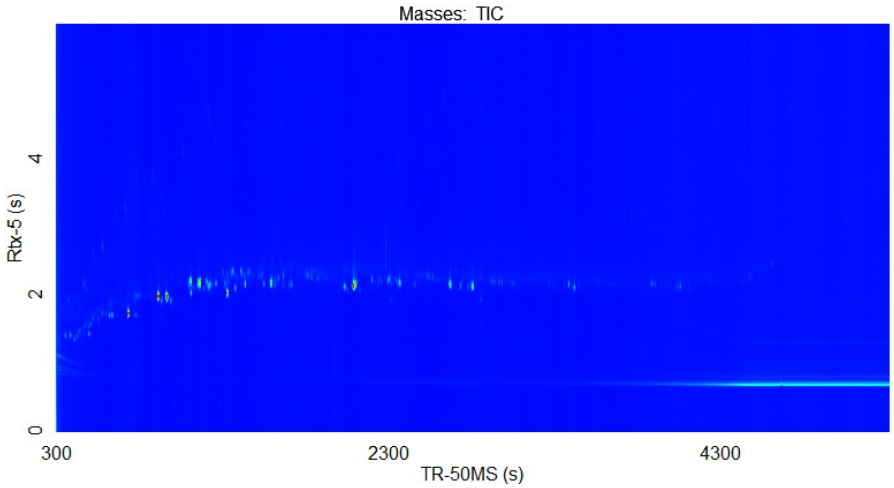
D4V: TIC



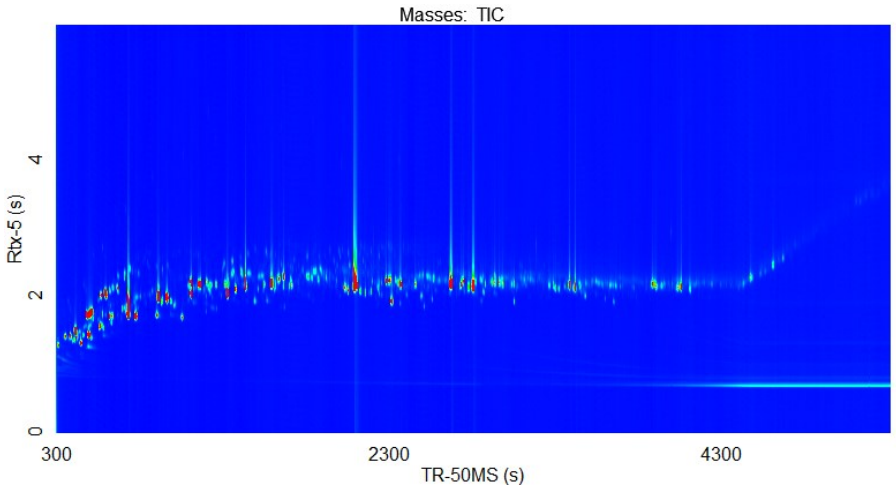
D5V: TIC



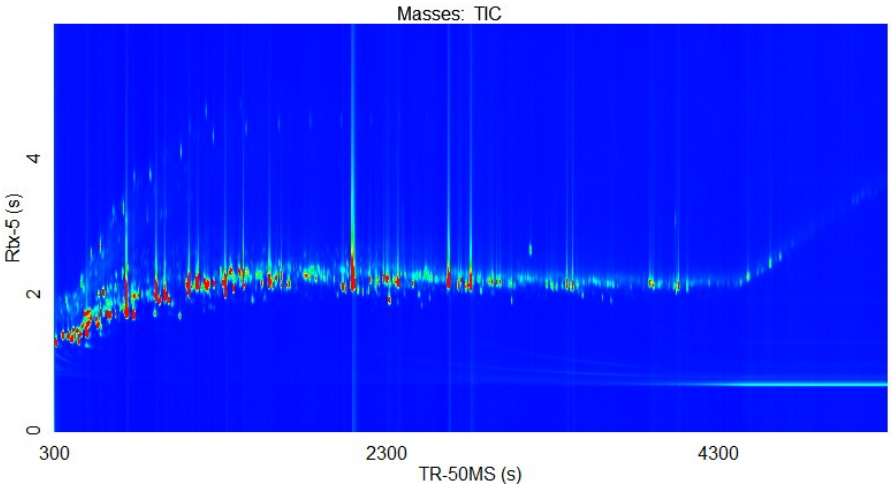
D6V: TIC



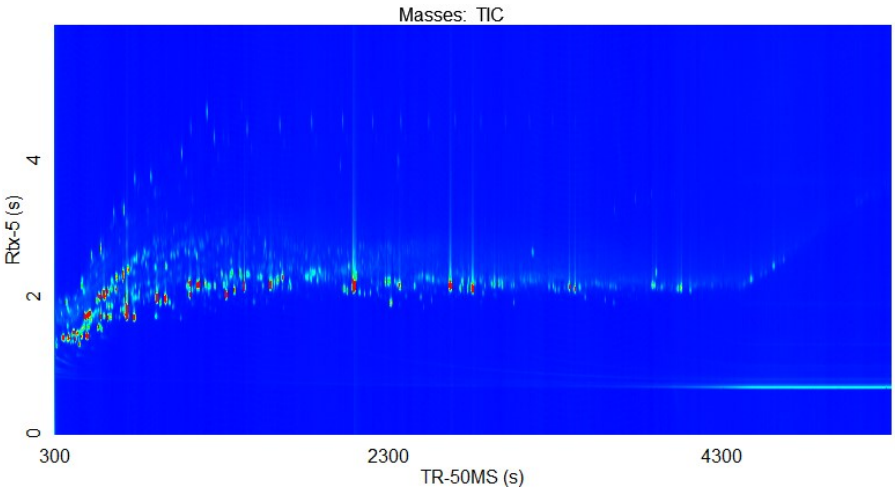
D7H: TIC



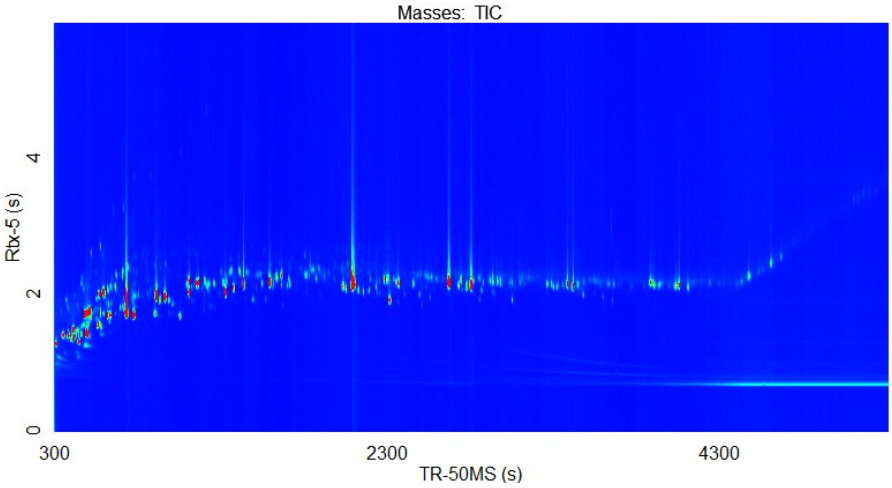
D8H: TIC



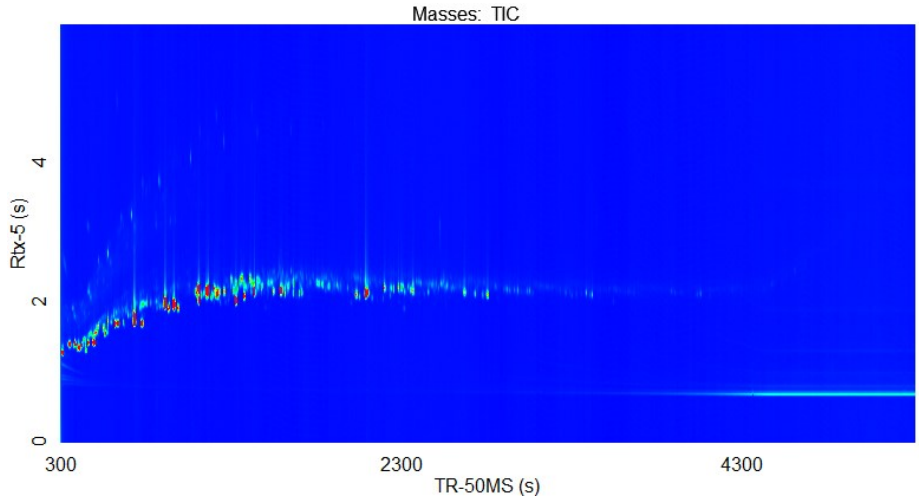
D9L: TIC



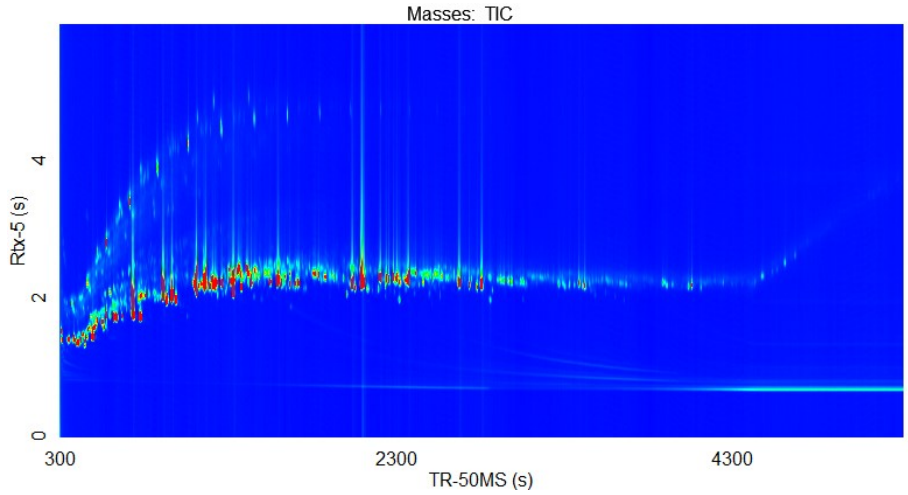
D10H: TIC



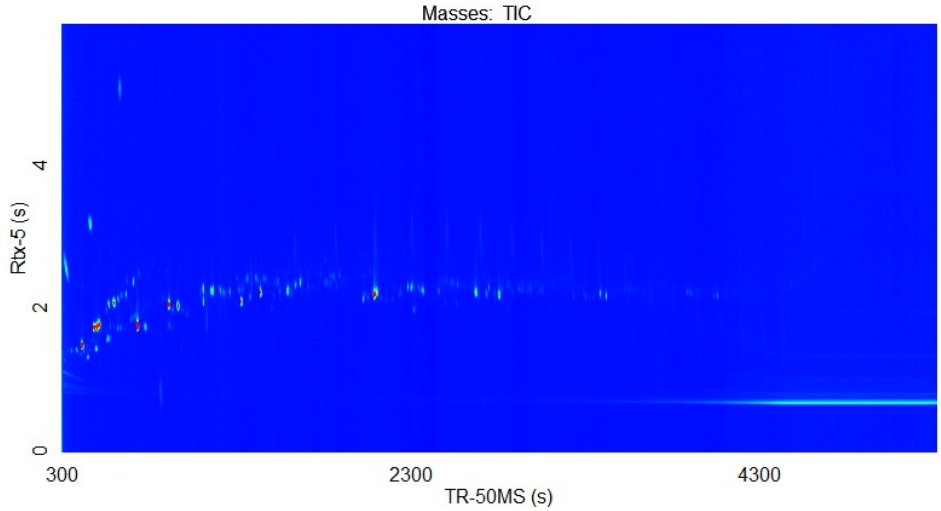
D13C: TIC



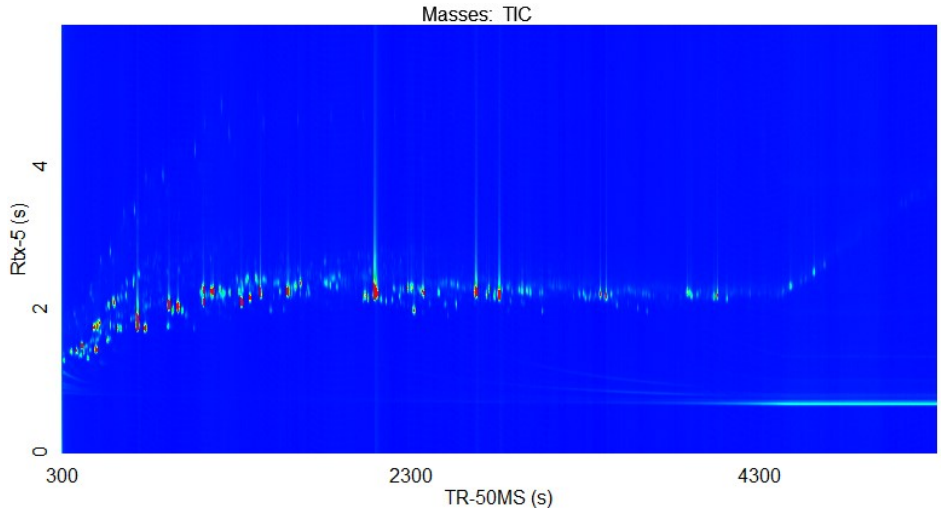
D14C: TIC



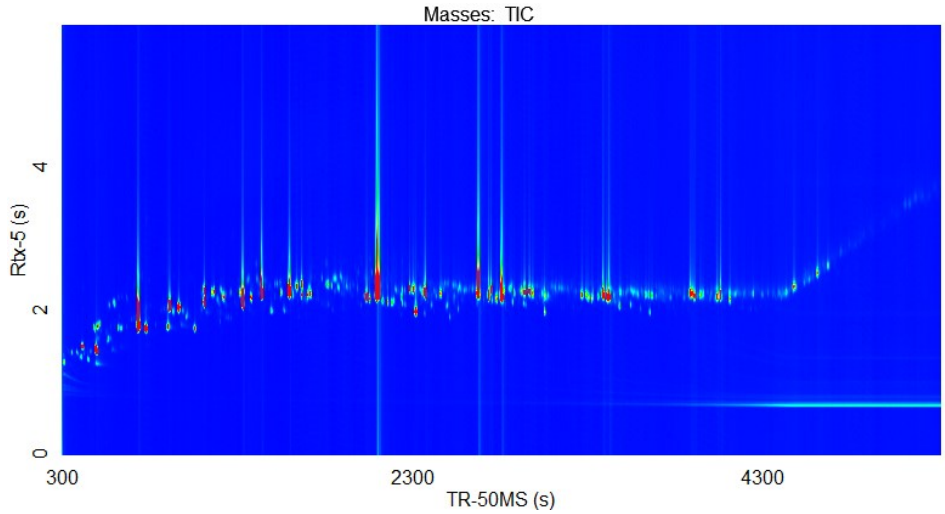
D16L: TIC



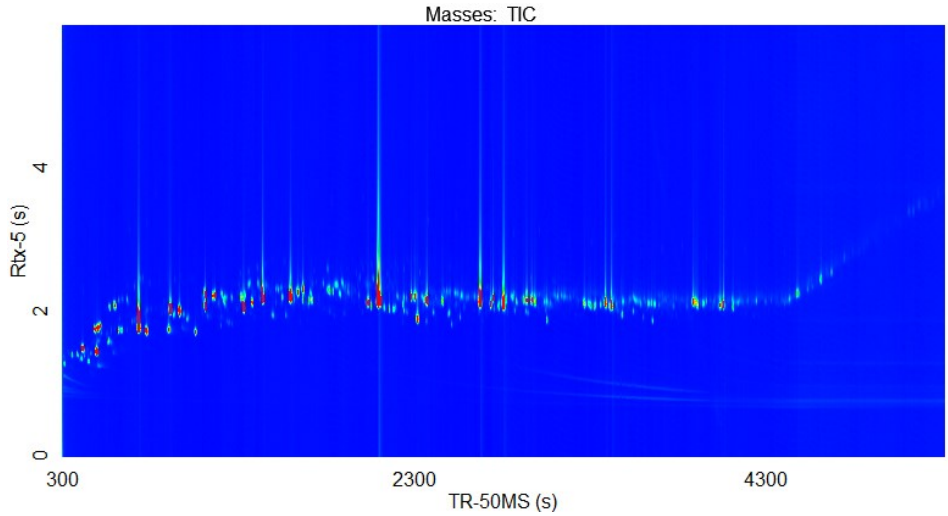
D17H: TIC



D18CO: TIC



D19CO: TIC



D20V: TIC